Parton Distribution Functions and Amplitudes of the Pseudoscalar Mesons and Nucleon from Lattice QCD

Bipasha Chakraborty Cambridge University Robert Edwards, Balint Joo, Jianwei Qiu, David Richards^{*}, Raza Sufian, Frank Winter Jefferson Lab Carl Carlson, Colin Egerer, Joe Karpie, Tanjib Khan, Kostas Orginos College of William and Mary Wayne Morris, Anatoly Radyushkin Old Dominion University Savvas Zafeiropoulos University of Heidelberg

March 4, 2019

Abstract

A long-standing challenge in lattice QCD is the direct computation of key measures of hadron structure, including parton distribution functions, quark distribution amplitudes, and three-dimensional measures such as the transverse-momentum-dependent distributions, and generalized parton distributions. Recently, new approaches have been proposed that enable their direct computation, and these are characterized by a requirement that the hadron of interest be increasingly relativistic. The aim of this **Class A Continuation** proposal is to capitalize on recent developments, and our previous USQCD allocation, to compute the structure functions of the pion and nucleon using the pseudo-PDF and current-current matrix element formulation using the distillation framework, and to investigate the principle systematic uncertainties, namely those arising from the finite volume of the lattices, and the finite discretisation and hence limits on the achievable spatial momentum. This project is relevant to the hadron structure experimental programs at JLab, RHIC-spin and at a future EIC, where lattice calculations are key to capitalization on the experimental efforts. We request an allocation of 89.2M KNL-core-hours, preferably on the cluster at JLab, and 900K GeForge GPU-Hours, 540 TByte of total tape storage (an additional 50), and 60 Tbyte of disk storage.

^{*}email: dgr@jlab.org

1 Goals and Milestones

Nuclear physics is a vast, rich field, whose phenomenology has been explored for decades through intense experimental and theoretical efforts. However, it is only now that direct connections to quantum chromodynamics (QCD), the underlying theory of the strong interactions, and the basic building blocks of nature, quarks and gluons, are solidifying. The phenomena of confinement and chiral symmetry breaking are jointly responsible for the intervening forty years of research and development between the discovery of QCD [1, 2, 3], the formulation of QCD on a discretized space-time [4] and recent numerical solutions of simple systems at the physical values of the quark masses along with the inclusion of electromagnetism. This new-found ability is changing nuclear physics. It is permitting us to address key questions that define the field, such as how the observed particles emerge from QCD, how gluons manifest themselves in the spectrum and structure of hadrons, and how the spin and other attributes of the pion and proton are partitioned between the quarks and gluons. The target for this proposal is to refine critical aspects of the nature of strongly interacting particles composed of quarks and gluons and their dynamics by direct calculations at the physical light quarks masses using the numerical technique of lattice QCD.

One of the great challenges posed by QCD is to understand how protons and neutrons, the basic building blocks of most of the observed matter in the universe, are made from quarks and glue. First principles calculations in this area are directly relevant to the experimental programs at JLab 12 GeV and LHC, and the future electron-ion collider (EIC). Our knowledge of hadron structure is encapsulated in a variety of measures. From the earliest observations of Bjorken scaling in Deep Inelastic Scattering, a one-dimensional longitudinal description of the nucleon has been provided through the unpolarized and polarized *Parton Distribution Functions* (PDFs). In contrast, the transverse distribution of charge and currents was probed in elastic scattering, encapsulated in the electric and magnetic form factors. More recently, new measures have been discovered correlating both the longitudinal fractional momentum x and impact-parameter space, and the *Transverse-Momentum-Dependent Distributions* (TMDs) providing a description in x and transverse momentum space. These new descriptions have opened new vistas on the nucleon, in particular enabling orbital angular momenta to be discerned.

From the inception of these new measures, the lattice community has attempted to calculate them from first principles. However, the formulation of lattice QCD in Euclidean space provides a formidable restriction: the x dependence could not be computed directly, but only the x moments of the distributions. Furthermore, the breaking of rotational symmetry on the lattice in practice restricted such calculations to only the first few moments. Recently, new ideas have been proposed that aim to circumvent one or more of these restrictions [5, 6, 7, 8, 9, 10, 11, 12].

The work proposed here will explore these new methods to investigate the properties of the pion, the lightest hadron composed of the light u, and d, and of the nucleon through calculation of so-called good "Lattice Cross Sections" identified in ref. [12], and through the calculation of pseudo-PDFs[11]. In particular, we will calculate the collinear one-dimensional x-dependent parton distribution function of the pion and nucleon, notably at large x. The calculations proposed here can also be used to compute the moments of the parton distributions for comparison with earlier works [13]. The determination of large-x part of the pion PDF is the goal of the approved experiment C12-15-006 at Jefferson Lab, whilst the determination of the kaon structure function is the goal of the approved run-group C12-15-006A. Further, we will calculate the so-called quark distribution amplitude of the pion, relevant both for understanding the form factors at large Q^2 , and for processes such as Deeply-Virtual Meson Production (DVMP) being explored at Jefferson Laboratory, and for the exclusive pion-nucleon Drell-Yan experiment proposed for J-PARC in Japan. Furthermore, the approved experiments E-03-012, E-00-002, at Jefferson Lab (12 GeV) are targeting the determination of the nucleon structure functions at large-x. The large-x region is precisely the region our computational methods are more effective and therefore our work is in close synergy with the already approved experimental efforts at Jefferson Lab. In addition to the experimental program at Jefferson Lab, at the Theory Center, the JAM collaboration aims towards the determination of PDFs from experimental data. Our future plans include a close collaboration with JAM using our lattice QCD results to further enhance the fidelity of the pion and nucleon PDFs.

Finally, the computationally expensive building blocks of our adopted approach can be employed to study the structure of the nucleon and other hadrons. More generally, this proposal will establish the theoretical and computational underpinnings for future calculations of the structure of hadrons.

The new measures of hadron structure entailed the computation of so-called pseudo-distributions [14, 11] and quasi-distributions [5], which are defined by Fourier transforms of the matrix element

$$M^{\alpha}(-P \cdot z, -z^2) = \langle P \mid \bar{\psi}(z)\gamma^{\alpha} \exp\left(-ig \int_0^z dz' A^z(z')\right)\psi(0) \mid P \rangle, \tag{1}$$

with respective to the *Ioffe Time* $\nu = -P \cdot z$ and the spatial separation z, respectively, where gauge invariance is implemented through a straight Wilson Line between $\psi(z)$ and $\bar{\psi}(0)$. Using $z = (0, 0, 0, z_3)$, $\alpha = 0$, and the hadron momentum $p = (p_0, 0, 0, p_3)$ one gets that the ratio

$$\mathfrak{M}(\nu, z_3^2) \equiv \frac{\mathcal{M}^0(\nu, z_3^2)}{\mathcal{M}^0(0, z_3^2)} , \qquad (2)$$

which is finite in the continuum limit and requires no renormalization, is directly related to PDFs. The PDFs can be extracted through a non-local operator product expansion where short distance contributions are computed perturbatively as discussed in [15]. In contrast to the matrix element defining the PDFs, the quark fields ψ and $\bar{\psi}$ are defined at equal time, and thereby admit calculation on a Euclidean lattice. At sufficiently short distances z_3 , $\mathfrak{M}(\nu, z_3^2)$ can be related to the familiar standard PDFs with calculable matching coefficients, and several recent works have computed these pseudo-quark distributions in ref. [9] (or the corresponding quasi-quark distributions in ref. [16, 17]), and attempted to relate them to the well-known standard distributions in the nucleon.

More recently, it has been shown that these various measures of lattice hadron structure could be interpreted in terms of a class of so-called "lattice cross sections" (LCSs)[12], computable directly in lattice QCD, that are factorizable into parton distribution functions with calculable coefficients, in the same manner as the hadronic cross sections measured in experiment. In particular, these lattice cross sections are expressed as single-particle matrix elements of non-local operators $\mathcal{O}_n(z)$:

$$\sigma_n(\nu, -xi^2) = \langle P \mid T\{\mathcal{O}_n(\xi)\} \mid P \rangle \tag{3}$$

where n labels the operator, P is the hadron momentum and ξ is the largest separation of the fields in the operator \mathcal{O}_n . These LCS can then be related to the parton distribution functions $f_a(x)$, where a labels the parton flavor, through

$$\sigma_n(\nu, -\xi^2) = \sum_a \int_{-1}^1 \frac{dx}{x} f_a(x, \mu^2) K_n^a(x\nu, z^2\mu^2) + \mathcal{O}(\xi^2 \Lambda_{\text{QCD}}^2), \qquad (4)$$

where μ is the factorization scale. The kernel K_n^a are calculable in (continuum) perturbation theory.

Here, we propose to employ gauge-invariant currents, separated in space, for example:

$$\mathcal{O}_{S}(\xi) = (\xi^{2})^{2} Z_{S}^{2} [\bar{\psi}_{q} \psi_{q}](\xi) [\bar{\psi}_{q} \psi](0) \mathcal{O}_{V'}(\xi) = \xi^{2} Z_{V'}^{2} [\bar{\psi}_{q}(\xi \cdot \gamma) \psi_{q'}](\xi) [\bar{\psi}_{q'} \xi \cdot \gamma \psi](0),$$
(5)

representing scalar and flavor-changing current combinations respectively, and Z_S , $Z_{V'}$ are the relevant quark bi-linear renormalization constants that render the left hand side of eq. 4 renormalization group invariant and absorb the UV divergences of the quark bi-linear currents. Note that the quark flavor q' in Eq. (5) is not required to correspond to a physical quark in the hadron, but can be a heavy "auxiliary" quark as we note below. The matrix element would become sensitive to the PDFs of both quarks, but the "auxiliary" quark distribution will be negligible in the valence distribution. A heavier quark improves the computational efficiency of the method, by effectively limiting the accessible phase space for propagation between the two currents. This procedure has been also suggested for improving the signal in lattice calculations of x moments of distributions in ref. [18].

The lightest pseudo-scalar mesons, the pion, composed of u and d quarks and antiquarks, and the kaon, containing a strange quark or antiquark, are particularly attractive computational theaters in which to explore these new ideas in hadron structure, as they are a simple hadronic state as well as playing an important role in low energy QCD dynamics. The structure of the pion is of inherent theoretical interest and plays a vital role in understanding the nucleon and in nuclear structure. The \bar{u} quark distribution of the π^- has been extracted from the FNAL E615 experiment [19] in a leading-order analysis. This extracted \bar{u} quark distribution has a tantalizing discrepancy with expectations from perturbative QCD (pQCD), and with various models of QCD. This is especially notable in its large-x behavior, which has different predictions in different theoretical models. Among the suggestions, it has been emphasized in Ref. [20] that the next-to-leading-logarithmic threshold re-summation effects in the calculation of the Drell-Yan cross section are important and give a softer valence distribution which falls of as $(1 - x)^2$ near $x \to 1$, consistent with the prediction based on the framework of perturbative QCD in Refs. [21, 22, 23]. There are also different model predictions for the large-x behavior of the pion valence distribution, some of which predict a harder fall-off as (1 - x) [24, 25, 26] or $(1 - x)^2$, such as in Dyson-Schwinger type models [27, 28].

Progress with Current allocation

Pion PDF with "Lattice Cross Section"

We have recently performed the first calculation of the pion valence PDF $q_V^{\pi}(x) = u(x) - \bar{d}(c)$ using the LCS approach[29]. The calculation was performed at a relatively large pion mass of around 400 MeV, and at a coarse lattice spacing of $a \simeq 0.127$ fm. The left-hand panel of Figure 1 shows the lattice cross section for the anti-symmetric "V-A" current combination as a function of the Ioffe time $\nu = p \cdot \xi$ and quark separation $z = (0, 0, \xi, 0)$, for the values of p_z admitted on our lattice, together with a fit to the data. Determining the parton distribution function from eq. 4 involves both a knowledge of the perturbative kernel represented by K_n^a and the solution of the "inverse" problem of eq. 4. We use the lowest-order (tree-level) calculation of the kernel, and address the inverse problem through the use of a phenomenologically motivated functions for form for the quark PDF, namely

$$q_{\rm v}^{\pi}(x) = \frac{x^{\alpha}(1-x)^{\beta}(1+\gamma x)}{B(\alpha+1,\beta+1) + \gamma B(\alpha+1+1,\beta+1)}.$$
(6)

A comparison of the resulting valence-quark PDF, and model- and phenomenological parametrizations is shown as the right-hand panel in Figure 1

Nucleon and Pion Pseudo-PDF

We are currently completing the calculation of the nucleon and pion PDF using the pseudo-PDF approach, in which the two currents are replaced by a quark and antiquark connected by a Wilson line, introduced in Equation 1. The analysis proceeds in an analogous manner to that described



Figure 1: Left-hand panel: Fit to the antisymmetric V-A currents matrix element with leading order (LO) perturbative kernel in Eq. 4 and functional form of pion valence distribution in Eq. 6. Right-hand panel: comparison of pion $xq_v^{\pi}(x)$ -distribution with the leading-order (LO) extraction from Drell-Yan data [19] (gray data points with uncertainties), next-to-leading order (NLO) fits [30, 20, 31] (orange band, magenta curve, and red band), and model calculations [26, 28] (black and blue lines). This lattice QCD calculation of $q_v^{\pi}(x)$ is evolved from an initial scale $\mu_0^2 = 1$ GeV² at LO. All the results are evolved to an evolution scale of $\mu^2 = 27$ GeV².



Figure 2: The left- and right-hand panels show the pion and nucleon parton distribution functions, respectively, including comparisons with phenomenological fits and QCD-inspired model calculations.

for the current-current correlators. Preliminary results for the extracted pion and nucleon parton distribution functions are shown as the left- and right-hand panels of Figure 2.

Nucleon Charges with Distillation

A detailed comparison of the calculation of nucleon charges using distillation, and the more usual method of Jacobi smearing, was performed in ref. [32], illustrated in Figure 3. The figure clearly demonstrates both the clear statistical advantage enjoyed by distillation over Jacobi smearing through the improved volume-sampling of the lattice, and the improved control over excited-state contamination through the use of the variational method. The use of distillation will be a central element of

the work proposed for this year.



Figure 3: The nucleon isovector scalar charge obtained using a standard Jacobi smearing algorithm (left), and distillation with a variational basis of seven operators (right), obtained on a $32^3 \times 64$ lattice with a pion mass around 360 MeV, and a lattice spacing $a \simeq 0.1$ fm.

Work for 2019-2020 Allocation Year

Impressive as the results presented above appear, they are still subject to as yet undetermined systematic uncertainties that must explored before the calculations can confront experiment, and add in a rigorous way to our knowledge of hadronic physics. The aim of the resources requested here is to compute x-dependent parton distribution functions of the pion and nucleon for a variety of "good" lattice cross sections, including the pseudo-PDFs for quark and anti-quark fields separated by a Wilson line, and for correlation functions of gauge-invariant spatially-separated current-current correlators. Capitalizing on the work described in ref. [29], we will work at a finer lattice spacing of $a \simeq 0.09$ fm, and at close-to-physical light-quark masses. This program will enable us to have fine control over the systematics of the calculation, and to seek signatures of higher-twist effects.

We note that the quark distribution amplitudes for pion involve an analogous set of currents and methodology to those of the parton distribution functions, except that they involve the matrix elements between the *vacuum* and the pion or kaon. An understanding of their behavior has long been sought, both because they are phenomenologically necessary for interpreting exclusive meson production and Drell-Yan experiments, and because they provide information about the approach to a partonic picture of hadron structure. There has been a recent effort at computing these distribution functions in lattice QCD in the current-current correlator approach [33], and we will also compute them here using both the pseudo-PDF approach, and using current-current matrix elements.

Finally, as is clear from a comparison of the pion PDFs in Figures 1 and 2, the LCS approach requires a larger ensemble of configurations than that for the pseudo-PDF approach to attain comparable statistical quality since it does not admit performing a full sampling of the lattice at the operator timeslice.



Figure 4: A schematic diagram of the hadron correlation function, for the case of the nucleon. The propagators joining the two currents J_1, J_2 can acquire any momentum depending on the available phase space for the case of the current-current correlators. For the case of the pseudo-distribution, the dashed line corresponds to the Wilson line of eqn. 1.

Computational Strategy

Our strategy will be to use an already generated gauge ensemble, with a clover-fermion action, a lattice spacing $a \simeq 0.09$ fm, $64^3 \times 128$ lattice sites, and close-to-physical u, d and s quark masses. This ensemble has the advantage that the matching coefficients, the Z of eqn. 5, have been previously computed in ref [34], thereby significantly reducing the costs of the computations proposed here. In contrast to the work of the last allocation period, we are adopting a different computational strategy of distillation, as we explain below.

The three-point function corresponding to hadronic matrix elements for two currents J_1 and J_2 separated in a Euclidean direction by ξ is shown in Figure 4, and can be written

$$C^{3\text{pt}}(t_f, t_i, t; x_0, \xi, \vec{p}) = \langle \mathcal{O}_p(\vec{z}, t_f) \mathcal{J}_1^{\dagger}(x_0 + \xi, t) \mathcal{J}_2(x_0, t) \mathcal{O}_p^{\dagger}(\vec{y}, 0) \rangle$$
(7)

where x_0 is a randomly determined source point, $\mathcal{J}_1, \mathcal{J}_2$ are currents of the form $q\Gamma Q$ where Q is an auxiliary quark field propagating on time slice t, and $\mathcal{O}\vec{p}$ is an interpolating operator for the hadron of study at momentum \vec{p} ; note that for the pseudo-PDF, the propagator corresponding to Q is simply a Wilson line. Because we are not performing a time slice momentum projection at the operator, the standard sequential-source method will not work here. However, for the case of mesons there is a straightforward variation of the sequential source method that was employed in ref. [29]. There is no analogous approach for baryons such as the nucleon, and we are exploiting a different approach, namely the use of "distillation", which both allows us to explore baryon structure, and offers other advantages, as we detail below.

Distillation

Distillation[35] is a variant of the "smearing" procedure. The use of "smearing" in LQCD calculations, in which the quark fields are replaced by extended objects through some gauge-invariant, spatially

symmetric operation:

$$\tilde{\psi}(\vec{x},t) = \sum_{\vec{y}} L(\vec{x},\vec{y})\psi(\vec{y},t)$$
(8)

has long been used as a means of improving the projection onto the low-lying states in correlation functions. Distillation and its descendants rely on the observation that we can write such as smearing matrix, for example that of three-dimensional Jacobi-smearing, as

$$L^{(J)} \equiv (1 - \frac{\kappa}{n} \Delta)^n = \sum_{i=1} f(\lambda_i) v^{(i)} \otimes v^{*(i)}, \qquad (9)$$

where $v^{(i)}$ is the *i*th (three-dimensional) eigenvector of Δ with eigenvalue λ_i , and $f(\lambda_i) = (1 - \frac{\kappa \lambda}{n})^n$. We now truncate the sum in eqn. 9 at some sufficient N_{vec} that the low-energy-scale physics is faithfully captured. We will take $f(\lambda) = 1$, and obtain the so-called distillation function

$$\Box(t) = \sum_{i=1}^{N_{\text{vec}}} v^{(i)}(t) v^{(i)\dagger}(t).$$
(10)

which we now use as the kernel for our quark smearing in eqn. 8. Focusing on the mesons, we write a "smeared" two-point correlation function as

$$C_{AB}(t_f, t_i; \vec{p}) = \text{Tr} \langle \Phi^A(\vec{p}, t_f) \tau(t_f, t_i) \Phi^B(\vec{p}, t_i) \tau^{\dagger}(t_f, t_i) \rangle,$$
(11)

where

$$\Phi_{\alpha\beta}^{A,ij}(\vec{p},t) = \sum_{\vec{x}} v^{*(i)}(\vec{x},t) e^{-i\vec{p}\cdot\vec{x}} [\Gamma^A(t)\gamma_5]_{\alpha\beta} v^{(j)}(\vec{x},t)$$
(12)

$$\tau_{\alpha\beta}^{ij}(t_f, t_i) = v^{*(i)}(t_f) M_{\alpha\beta}^{-1}(t_f, t_i) v^{(j)}(t_i).$$
(13)

Here $i, j = 1, ..., N_{\text{vec}}, \alpha, \beta$ are spinor indices, and t_i, t_f the source and sink timeslices respectively, and Γ is an operator that may include derivatives.

The crucial observation in the above construction is that the Φ 's of eqn. 12 encode the structure of the operator, with, for example, derivatives now acting on the (quark-mass-independent) ξ 's, whilst the parallel transport of the quark fields is encoded in the "perambulators" of eqn. 13. This factorization enables operators to be constructed not only at the sink but also at the source with essentially unrestricted spatial structures, for any set of perambulators.

The calculation of three-point functions is discussed in ref. [36]; crucially, it involves the introduction of so-called "generalized perambulators",

$$S^{ij}(t_f, t_i, t; x_0, \xi) = v^{(i)\dagger}(t_f) M^{-1}(t_f; x_0 + \xi, t) \Gamma_1 G_Q(x_0 + \xi; x_0, t) \Gamma_2 M^{-1}(x_0, t; t_i) v^{(j)}(t_i)$$
(14)

where G_Q is the spectator quark propagator corresponding to the eqn. 7. involves the quark and anti-quark fields at some displacement, for example corresponding to the operators of eqn 1. These generalized perambulators are trivially constructed from the solution vectors and the three-point function may be written

$$C_{AB}^{3\text{pt}}(t_f, t_i, t; x_0, \xi, \vec{p}) = \text{Tr}\langle \phi^A(\vec{p}, t_f) S(t_f, t_i, t; x_0, \xi) \Phi^B(\vec{p}, t) \tau^{\dagger}(t_f, t_i) \rangle,$$
(15)

The construction of the generalized perambulators required for quark distribution amplitudes is the same as that of eqn 14, with the restriction that t_f and t_i are now the same. The resulting correlation functions are now two-point functions. Note that the evaluation of eqn. 15 is identical to that of

	Task	$N_{\rm eigen}$	N_{src}	$N_{\rm sep}$	$N_{ m mom}$	$N_{ m inv}$	$N_{\rm cfg}$	TOTAL (KNL core-hours)
(1)	Soln	64	8	8	2	32,768	350	$65.4\mathrm{M}$
(2)	G_Q		8	8		$6,\!144$	350	$12.3\mathrm{M}$
(3)	Genprop	64	8	8	2		350	$1.7\mathrm{M}$
(4)	Elem	64			4		700	$9.8\mathrm{M}$
TOTAL								89.2M

Table 1: The table shows the cost in KNL core-hours for each task, as described in the text. The total requested time is is 89.2M KNL-core-hours.

eqn. 11, except that on of the perambulators is replaced by the generalized perambulator, and the possible additional Wick contractions associated with different possible insertions of the current.

The construction of two- and three-point functions for baryons proceeds likewise:

$$C_{AB}(t_{f}, t_{i}; \vec{p}) \simeq \langle \phi^{A}_{i'j'k'}(\vec{p}, t_{f})\tau^{i'i}(t_{f}, t_{i})\tau^{j'j}(t_{f}, t_{i})\tau^{k'k}(t_{f}, t_{i})\phi^{B}_{ijk}(\vec{p}, t)\rangle$$

$$C^{3\text{pt}}_{AB}(t_{f}, t_{i}, t; x_{0}, \xi, \vec{p}) \simeq \langle \phi^{A}_{i'j'k'}(\vec{p}, t_{f})\tau^{i'i}(t_{f}, t_{i})\tau^{j'j}(t_{f}, t_{i})S^{k'k}(t_{f}, t_{i})\phi^{B}_{ijk}(\vec{p}, t_{i})\rangle$$
(16)

up to additional Wick contractions.

Distillation affords several distinct advantages for the work proposed here. Firstly, the factorization of the quark propagation and the hadronic wave function described above ensures that the generalized perambulators of eqn. 14 are dependent only on the operator inserted at timeslice t, and not on the interpolating operator at source and sink, and thus the same relatively expensive, but compact, generalized perambulators can be used for investigations of hadron structure for any external state, including both pion and nucleon. Secondly, the method enables a momentum projection to be performed at both the source and sink interpolating operators and, furthermore, for the case of the pseudo-PDF, at the current operator, thus ensuring a better sampling of the gauge configurations. Thirdly, the method enables the variational method to be applied using a broad variational basis without the computation of additional quark propagators. Finally, the calculation of current-current matrix element for the case of baryons simply cannot be accomplished using the extension of the sequential-source method that we used for the pion.

Software

We will use the *Chroma* software framework, which now supports multigrid on both the GPUs and on the KNLs. The elementals are computed using the *harom*, running on top of the three-dimensional qdp++ code.

Resource Request

We are requesting resources on both the KNL and GeForce GPUs; we will use the former for both in the pseudo-PDF and LCS calculations, and the latter to achieve the necessary statistical precision in the LCS approach. The computational request encompasses four tasks. Task (1) is the calculation of the solution vectors. Task (2) is the calculation of the spectator quark propagators in the LCS calculation, corresponding to the propagation of the spectator quark Q between the two currents \mathcal{J}_1 and \mathcal{J}_2 of Figure 4. Task (3) is the contraction of the solution vectors to form the generalized perambulators for the calculation of the pseudo-PDFs. Task (4) is the calculation of the so-called elementals encoding the wave functions for baryons. For the case of the calculation of LCS, we do not

	Task	$N_{\rm eigen}$	N_{src}	$N_{\rm sep}$	$N_{\rm mom}$	$N_{ m inv}$	$N_{\rm cfg}$	TOTAL (gpu-hours)
(1)	Soln	64	8	8	2	32,768	350	757k
(2)	G_Q		8	8		6,144	350	142k
TOTAL								900K

Table 2: The table shows the cost in GeForce Node-Hours for each task, as described in the text. The total requested time is is 900K GPU node-hours.

need to perform the construction of the generalized perambulators in memory since we will write out a small, reduced propagator for both the spectator and active quarks which can then be contracted off-line. We emphasize that we are able to exploit both the KNLs and GPUs for Tasks (1) and (2). Tasks (3) and (4) can be performed on the KNLs The timings for each of these tasks is given in Table 1 and Table 2, as we describe below.

Our timings for the calculation of the solution vectors are obtained from the measured time on the target lattice obtained on the KNL nodes at Jefferson Lab, for which we find the multigrid solver requires 5.7 KNL core-hours when using a 32 KNL (16p/18p) nodes. Whilst a single solve can be performed on fewer than 32 nodes, a key part of our strategy for the calculation of pseudo-PDFs is to keep the computed solution vectors in memory to avoid expensive input/output to disk. We will use a $N_{\rm vec} = 64$ eigenvectors, but with $N_{\rm mom} = 2$ momentum boosts. We will compute the three-point functions from a total of $N_{\rm src} = 8$ source points on each lattice, and for a total of $N_{\rm sep} = 8$ source-sink separations. Thus the total number of inversions per configuration is $N_{\rm inv} = 4N_{\rm vec}N_{\rm mom}N_{\rm src}N_{\rm sep} =$ 32768, where the factor of four reflects the number of spinor indices. For the GPU, we obtain our timings from the observed 0.022 GPU-hours needed for a single solve on the target lattice using the Volta GPUs on Summit; we assume a calculation on the GeForce GPUs will require three times the GPU-hours, ie 0.066 GPU-hours.

The second task is the solution of the spectator quark propagators between the two currents. This will be performed for each combination of the source, sink and current-insertion time slices, yielding a total of $8 \times 8 \times 8 \times 12$ inversions, where the final factor of 12 denotes the number of spin-color sources corresponding to a quark propagator; we assume that the solution time will be the same as that quoted above.

The third task is the calculation of the generalized perambulators needed for the calculation of the pseudo-PDFs, which we will perform on the KNLs in the same job as that for tasks (1) and (2) above; note that the lack of ECC precludes the GeForce GPUs from directly performing this task. Our time estimate is based on the measured time to calculate a single generalized perambulator at a single timeslice, obtained on a 32^3 lattice with $N_{\text{vec}} = 64$ and $N_{\text{mom}} = 1$, which requires 2.5 seconds on a single dual-socket eight-core Sandy-Bridge (12s) node, corresponding to 20.71 KNL core-hours, using the 2018 USQCD conversions factors. Thus the time to compute a single generalized perambulator at a single time slice for the target lattice is $(64/32)^3 \times 20.71 = 166$ KNLcore – seconds. For our calculation here, we will calculate the generalized perambulators for $N_z = 25$ separations between the currents, and for four γ matrices.

The final task is the construction of the elementals. The timings here are based on the observed performance for a single, local interpolating operator and two momenta on a single $32^3 \times 64$ lattice on a single KNL, which required one KNL node-hour. Scaling to 7 combinations of two derivative operators, and 4 momenta, we find a total cost per configuration of 14K KNL core-hours. This can only be done on the KNLs and therefore we will need to perform this task both for configurations to be analyzed on the KNLs and those on the GPUs.

We are currently using 490 TByte of tape storage. The largest increase in storage will arise from

the additional elementals of Table 1, which will require 30 Tbyte. Together with the perambulators, generalized perambulators, and correlators, we request an additional 50 Tbyte of tape storage, yielding 540 TByte, and 60 TByte of disk space, corresponding to 700 and 450 K Sky-core-hour respectively.

Readiness and Run Schedule

The software used in this proposal is that of USQCD, and in particular *Chroma*, and the codes are ready for production running. The GeForce GPUs lack ECC memory, and we will investigate workflow whereby the GPUs can be used for solves and the KNLs for contractions both for the pseudo-PDF and LCS approaches. Further, we would like to optimize some of our single-node contraction code for the KNLs.

Data sharing and exclusive rights

We note that most of the data generated as part of this proposal comprises the correlators rather than propagators and gauge configurations. There are in addition "perambulators" and solution vectors that are being used here for investigations that are being used here for studies of the excited-state nucleon spectrum and for reducing the excited-state contamination of nucleon matrix elements that could also be used for other projects and will be available for such.

References

- [1] Fritzsch H, Gell-Mann M and Leutwyler H 1973 Phys. Lett. B47 365–368
- [2] Politzer H D 1973 Phys. Rev. Lett. **30** 1346–1349
- [3] Gross D J and Wilczek F 1973 Phys. Rev. Lett. 30 1343–1346
- [4] Wilson K G 1974 Phys. Rev. **D10** 2445–2459
- [5] Ji X 2013 Phys. Rev. Lett. **110** 262002 (Preprint 1305.1539)
- [6] Ji X 2014 Sci. China Phys. Mech. Astron. 57 1407–1412 (Preprint 1404.6680)
- [7] Ma Y Q and Qiu J W 2014 (*Preprint* 1404.6860)
- [8] Liu K F 2016 PoS LATTICE2015 115 (Preprint 1603.07352)
- [9] Orginos K, Radyushkin A, Karpie J and Zafeiropoulos S 2017 Phys. Rev. D96 094503 (Preprint 1706.05373)
- [10] Chambers A J, Horsley R, Nakamura Y, Perlt H, Rakow P E L, Schierholz G, Schiller A, Somfleth K, Young R D and Zanotti J M 2017 Phys. Rev. Lett. 118 242001 (Preprint 1703.01153)
- [11] Radyushkin A V 2017 Phys. Rev. D96 034025 (Preprint 1705.01488)
- [12] Ma Y Q and Qiu J W 2018 Phys. Rev. Lett. **120** 022003 (Preprint 1709.03018)
- [13] Karpie J, Orginos K and Zafeiropoulos S 2018 JHEP 11 178 (Preprint 1807.10933)

- [14] Braun V, Gornicki P and Mankiewicz L 1995 Phys. Rev. D51 6036–6051 (Preprint hep-ph/ 9410318)
- [15] Radyushkin A 2018 Phys. Rev. D98 014019 (Preprint 1801.02427)
- [16] Lin H W, Chen J W, Cohen S D and Ji X 2015 Phys. Rev. D91 054510 (Preprint 1402.1462)
- [17] Alexandrou C, Cichy K, Drach V, Garcia-Ramos E, Hadjiyiannakou K, Jansen K, Steffens F and Wiese C 2015 Phys. Rev. D92 014502 (Preprint 1504.07455)
- [18] Detmold W and Lin C J D 2006 Phys. Rev. D73 014501 (Preprint hep-lat/0507007)
- [19] Conway J S et al. 1989 Phys. Rev. **D39** 92–122
- [20] Aicher M, Schafer A and Vogelsang W 2010 Phys. Rev. Lett. 105 252003 (Preprint 1009.2481)
- [21] Farrar G R and Jackson D R 1979 Phys. Rev. Lett. 43 246
- [22] Berger E L and Brodsky S J 1979 Phys. Rev. Lett. 42 940–944
- [23] Brodsky S J, Burkardt M and Schmidt I 1995 Nucl. Phys. B441 197–214 (Preprint hep-ph/ 9401328)
- [24] Shigetani T, Suzuki K and Toki H 1993 Phys. Lett. B308 383-388 (Preprint hep-ph/9402286)
- [25] Melnitchouk W 2003 Eur. Phys. J. A17 223–234 (Preprint hep-ph/0208258)
- [26] de Teramond G F, Liu T, Sufian R S, Dosch H G, Brodsky S J and Deur A (HLFHS) 2018 Phys. Rev. Lett. 120 182001 (Preprint 1801.09154)
- [27] Hecht M B, Roberts C D and Schmidt S M 2001 Phys. Rev. C63 025213 (Preprint nucl-th/ 0008049)
- [28] Chen C, Chang L, Roberts C D, Wan S and Zong H S 2016 Phys. Rev. D93 074021 (Preprint 1602.01502)
- [29] Sufian R S, Karpie J, Egerer C, Orginos K, Qiu J W and Richards D G 2019 (Preprint 1901. 03921)
- [30] Wijesooriya K, Reimer P E and Holt R J 2005 Phys. Rev. C72 065203 (Preprint nucl-ex/ 0509012)
- [31] Barry P C, Sato N, Melnitchouk W and Ji C R 2018 Phys. Rev. Lett. 121 152001 (Preprint 1804.01965)
- [32] Egerer C, Richards D and Winter F 2018 Phys. Rev. D (in press) (Preprint 1810.09991)
- [33] Bali G S, Braun V M, Gläßle B, Göckeler M, Gruber M, Hutzler F, Korcyl P, Schäfer A, Wein P and Zhang J H 2018 Phys. Rev. D98 094507 (Preprint 1807.06671)
- [34] Yoon B et al. 2017 Phys. Rev. D95 074508 (Preprint 1611.07452)
- [35] Peardon M et al. 2009 Phys. Rev. **D80** 054506 (Preprint 0905.2160)
- [36] Shultz C J, Dudek J J and Edwards R G 2015 (Preprint 1501.07457)